

Opening the Vaults of Academe

George S. Porter
Hema Ramachandran
Caltech

ASEE National Conference, Engineering Libraries Division

11-16 June 2005

Primary Records of Intellectual Productivity?

- Theses & Dissertations
- Technical Reports
- Working Papers
- Preprints
- Journal Articles

Electronic Theses & Dissertations

- Clearly originating from an institution
- Commonly collected by libraries
- Rules and regulations do not require negotiation beyond the local level

Rolling Start ...

- Optional deposit initially
- Proof of concept – It can be done!
- Work out the kinks with volunteers
- Early adopters tend to be tech-savvy
- Explore, learn, document

Look Back & Assess

- Evaluate rate of deposit
- Gauge workload – sys admin, liaison librarian, incremental load of students
- Numbers – has the initial collection increased visibility of dissertations/institution?

Mandatory ETD?

- Across the board mandate, if campus is sufficiently tech-savvy
- Staged introduction? Where to start?
- Disciplines likely to be tech competent/confident
 - Engineering
 - Physics
 - Mathematics

Gray Literature

- Long tradition of technical report series in engineering, especially Computer Science
- Gray lit provides extensive data, procedures, proofs, likely to be omitted due to space constraints in formal publications

Gray Literature – cont'd

- Working papers – common in Economics
- Preprints – Math, Physics, Astrophysics
- Engineering does not have exclusive dominion in gray literature.

Gray Lit – cont'd

- NCSTRL
 - Have a CS dept?
 - Do they participate?
 - Is the library involved?
 - How complete is online access?

I'm from the library. I'm here to help.

Gray Lit – cont'd

- Major laboratories/Institutes on campus
 - Likely to produce (have produced) tech report series.
 - Historically spotty library collections
 - Library may have shied away from Gray Lit
 - Lab/Institute may not have thought of letting the library know about the series

Journal Articles

There is a wealth of publisher formatted,
final version articles available for
harvesting.

The trick is to know where to look.

Journal Articles – cont'd

- Open Access journals
 - BioMed Central titles
 - PNAS (perhaps a 6 month wait)
 - Optics Express
 - Nucleic Acids Research
 - European Geosciences Union
(Copernicus.org)
- Walker-Prosser – OA by/buy the article

Journal Articles – cont'd

- Conspicuously cooperative publishers request/require acknowledgment
 - Annual Reviews
 - Cambridge University Press
 - Institute of Physics
 - IEEE
 - Nature Publishing Group
 - SIAM

Time – the 4th Dimension

- Scholarly publishing did not begin within our lifetimes, regardless of what my children claim about dinosaurs, the invention of colors, etc.

Too Much Theory?

- Let's see what it looks like in practice

Caltech Collection of Open Digital Archives (CODA)

- **Coda:** *Mus.* A passage of more or less independent character introduced after the completion of the essential parts of a movement, so as to form a more definite and satisfactory conclusion. -- Oxford English Dictionary

<http://coda.caltech.edu/>

Caltech's Electronic Thesis Database

- Dissertations <http://etd.caltech.edu/>
- Student instructions <http://library.caltech.edu/etd/>
- 1962 dissertations online
- “Your dissertation will be in very good company...”
 - Linus Pauling, Paul MacCready, Donald Knuth, Benoit Mandelbrot, Frank Oppenheimer, Maurice Biot

ETDs – going back in time

- Where to start digitizing?
 - Faculty
 - Distinguished Alumni
 - Nobel Laureates
 - Fellows of NAS, NAE, AAAS

Permissions -- ← 1986 →

Leverage through NDLTD

Technical Reports & Working Papers

- Did I mention NCSTRL?
- Robotics
- Earthquake Engineering
- Aero-/Hydrodynamics or Fluid Mechanics
- Electrical Engineering
- High Energy Physics (SPIRES | arXiv)

At a guess, university libraries rarely have 50% of tech report series produced on campus.

Offer a Partnership

- PARADISE – Parallel and Distributed Systems Group
<http://caltechparadise.library.caltech.edu/>
- Center for Advanced Computing Research
<http://cacr.library.caltech.edu/>
- Caltech Earthquake Engineering Research Laboratory Technical Reports
<http://caltecheerl.library.caltech.edu/>

Publishing Service

- Caltech Large-Eddy Simulation and Subgrid-Scale Modeling for Turbulent Mixing and Reactive Flows
<http://caltechlessgs.library.caltech.edu/>
- Fourth International Symposium on Cavitation. Hosted by Caltech, June, 2001
<http://cav2001.library.caltech.edu/>
- Caltech Oral Histories
<http://oralhistories.library.caltech.edu/>

Recruitment of Archives

- Here's the harder part – convincing faculty of the value of depositing journal articles, conference papers, and book chapters.
- Classic “The Chicken & the Egg Problem”
 - Show me it's worth time & effort
 - Results require critical mass x time

Not the Young & the Restless

Our early adopters are not the oft envisioned young visionaries, risking their tenure decisions on innovative/dicey publishing models.

Caltech's early adopters are mature faculty establishing comprehensive collections of decades of research output.

Who are the early adopters?

- Christopher E. Brennen
 - Richard L. and Dorothy M. Hayman Professor of Mechanical Engineering



- Theodore Y.-T. Wu
 - Professor of Engineering Science, Emeritus, 1996-.



Early Adopters – cont'd

- Wolfgang G. Knauss
 - Theodore von Kármán
Professor of Aeronautics
and Applied Mechanics,
Emeritus
- Thomas Kirk Caughey, 1927-2004
 - Richard L. and Dorothy M.
Hayman Professor of
Mechanical Engineering, Emeritus

Early Adopters – cont'd

- Ronald F. Scott
 - Richard L. and Dorothy M. Hayman Professor of Mechanical Engineering, Emeritus

Sometimes you need to make things happen

What is the Open Access publishing output
of your campus?

Download those articles --- NOW!

Harvest from OA sources.

Let's get going

**It's easier to ask forgiveness than
it is to get permission.**

Adm. Grace Hopper (1906-1992)

Permission & Forgiveness

- In early May 2005, CaltechAUTHORS had just over 200 items. All posted items had documented permissions from both publisher and author.
- Expanding the net – adding OA articles and other articles from CC* publishers, currently 353 and growing rapidly.

Caltech's #1 OA author

Amnon Yariv,

Martin and Eileen

Summerfield Professor of
Applied Physics and
Professor of Electrical
Engineering



75 years young, 4 postdocs, 6 grad students

How now, PNAS?

- Check the Open Access publishing option elections of your faculty
- Easiest route? Via PubMed Central
<http://pubmedcentral.gov/tocrender.fcgi?iid=13407>,
then 'Find in Page' for likely faculty

What about those tech reports?

Here's a significant clue –

If the reports were deposited with the library, it was done so with the intent to document the existence of the reports and to promote distribution of the information.

Software?

- Pick something.
- The biggest challenges in the field of digital collections are sociological, not technological.
- Changing platforms (Dienst → Eprints→?; ETD-DB → Eprints or D-Space or ?) is not a critical issue.

Software – cont'd

- Without content, who cares what software isn't used to not make it available?

Questions?

- Attempted Answers